



# Impact of a computed tomography-based artificial intelligence software on radiologists' workflow for detecting acute intracranial hemorrhage

Jimin Kim<sup>1</sup>  
 Jinhee Jang<sup>2,3</sup>  
 Se Won Oh<sup>1</sup>  
 Ha Young Lee<sup>1</sup>  
 Eun Jeong Min<sup>4</sup>  
 Jin Wook Choi<sup>5</sup>  
 Kook-Jin Ahn<sup>2</sup>

<sup>1</sup>The Catholic University of Korea, Eunpyeong St. Mary's Hospital College of Medicine, Department of Radiology, Seoul, Korea

<sup>2</sup>Seoul St. Mary's Hospital College of Medicine, The Catholic University of Korea, Department of Radiology, Seoul, Korea

<sup>3</sup>Applied Artificial Intelligence Research (A<sup>2</sup>IR), Institute for Precision Health (IPH), University of California, Irvine, USA

<sup>4</sup>The Catholic University of Korea College of Medicine, Department of Medical Life Sciences, Seoul, Korea

<sup>5</sup>Ajou University School of Medicine, Ajou University Hospital, Department of Radiology, Suwon, Korea

Corresponding author: Jinhee Jang, Kook-Jin Ahn

E-mail: znee@catholic.ac.kr, ahn-kj@catholic.ac.kr

Received 14 February 2025; revision requested 05 April 2025; accepted 05 May 2025.



Epub: 07.07.2025

Publication date:

DOI: 10.4274/dir.2025.253301

## PURPOSE

To assess the impact of a commercially available computed tomography (CT)-based artificial intelligence (AI) software for detecting acute intracranial hemorrhage (AIH) on radiologists' diagnostic performance and workflow in a real-world clinical setting.

## METHODS

This retrospective study included a total of 956 non-contrast brain CT scans obtained over a 70-day period, interpreted independently by 2 board-certified general radiologists. Of these, 541 scans were interpreted during the initial 35 days before the implementation of AI software, and the remaining 415 scans were interpreted during the subsequent 35 days, with reference to AIH probability scores generated by the software. To assess the software's impact on radiologists' performance in detecting AIH, performance before and after implementation was compared. Additionally, to evaluate the software's effect on radiologists' workflow, Kendall's Tau was used to assess the correlation between the daily chronological order of CT scans and the radiologists' reading order before and after implementation. The early diagnosis rate for AIH (defined as the proportion of AIH cases read within the first quartile by radiologists) and the median reading order of AIH cases were also compared before and after implementation.

## RESULTS

A total of 956 initial CT scans from 956 patients [mean age: 63.14 ± 18.41 years; male patients: 447 (47%)] were included. There were no significant differences in accuracy [from 0.99 (95% confidence interval: 0.99–1.00) to 0.99 (0.98–1.00),  $P = 0.343$ ], sensitivity [from 1.00 (0.99–1.00) to 1.00 (0.99–1.00),  $P = 0.859$ ], or specificity [from 1.00 (0.99–1.00) to 0.99 (0.97–1.00),  $P = 0.252$ ] following the implementation of the AI software. However, the daily correlation between the chronological order of CT scans and the radiologists' reading order significantly decreased [Kendall's Tau, from 0.61 (0.48–0.73) to 0.01 (0.00–0.26),  $P < 0.001$ ]. Additionally, the early diagnosis rate significantly increased [from 0.49 (0.34–0.63) to 0.76 (0.60–0.93),  $P = 0.013$ ], and the daily median reading order of AIH cases significantly decreased [from 7.25 (Q1–Q3: 3–10.75) to 1.5 (1–3),  $P < 0.001$ ] after the implementation.

## CONCLUSION

After the implementation of CT-based AI software for detecting AIH, the radiologists' daily reading order was considerably reprioritized to allow more rapid interpretation of AIH cases without compromising diagnostic performance in a real-world clinical setting.

## CLINICAL SIGNIFICANCE

With the increasing number of CT scans and the growing burden on radiologists, optimizing the workflow for diagnosing AIH through CT-based AI software integration may enhance the prompt and efficient treatment of patients with AIH.

## KEYWORDS

Acute intracranial hemorrhage, computed tomography, deep learning, artificial intelligence, radiologist, workflow, accuracy

Early and accurate detection of acute intracranial hemorrhage (AIH) on brain computed tomography (CT) is imperative due to the serious risks posed by this condition.<sup>1-3</sup> Timely diagnosis allows for immediate, life-saving intervention, whereas delayed detection can result in severe brain damage or death.<sup>2-4</sup> However, the rapidly increasing number of CT scans performed daily has placed a substantial burden on medical staff, including radiologists, potentially compromising the accuracy and timeliness of AIH diagnosis.<sup>5,6</sup>

In addition to the increasing workload, radiologists often face interruptions in their workflow due to various factors, such as urgent consultations, training of junior staff, and technical issues with imaging equipment.<sup>7-9</sup> These disruptions can lead to delays in image interpretation, increased cognitive load, and even diagnostic errors, particularly in high-stakes conditions such as AIH.<sup>10,11</sup> Such challenges underscore the importance of optimizing radiologists' workflow to ensure timely and accurate diagnoses.<sup>12</sup>

Recently, artificial intelligence (AI) has become a major focus in the field of neuroradiology, and numerous commercially available AI-based software programs have been developed for detecting acute cerebral findings.<sup>13-18</sup> Although previous studies have demonstrated the impressive stand-alone performance of these AI algorithms in diagnosing AIH and other stroke-related conditions on CT scans, their potential benefits for workflow optimization remain underexplored. Although early and prompt decision-making in AIH cases is critical for patient outcomes,<sup>2-4</sup> radiologists have traditionally relied on ambiguous prioritization systems such as stat, routine, or first-in, first-out (FIFO). This is largely because they are

unable to assess the urgency of each exam in the worklist before opening it in the picture archiving and communication system (PACS).<sup>19,20</sup> To address this issue, some studies have shown that integrating AI algorithms into the PACS can greatly improve turnaround time (TAT) by prioritizing images based on urgency, thereby facilitating faster intervention and improved outcomes.<sup>20-24</sup> Therefore, evaluating the impact of AI software on radiologists' workflow in real-world settings is crucial for advancing its practical integration.

This observational study aims to explore the impact of a commercially available CT-based AI software for detecting AIH on radiologists' diagnostic performance and their workflow in a real-world clinical setting.

## Methods

The retrospective study was performed in line with the principles of the Declaration of Helsinki and approved by the Eunpyeong St. Mary's Hospital's Institutional Review Board (protocol number: PC24RASI0078, date: June 2024), and informed consent was waived according to the decision of the board committee.

### Sample eligibility

A total of 1,375 non-contrast brain CT scans from patients with suspected AIH (including subdural, epidural, subarachnoid, intraparenchymal, and intraventricular hemorrhages) were potentially eligible over a 70-day period between December 1, 2023, and February 9, 2024. During this period, scans

were included based on the following criteria: (1) the first CT scan performed during the patient's clinical course, (2) acceptable image quality for interpretation, and (3) availability of complete radiologist reports. All potentially eligible CT scans were reviewed by a board-certified neuroradiologist with 11 years of experience (J.K.) according to these criteria. After review, 273 follow-up scans, 140 scans with major metal artifacts caused by clips or coils, and 6 scans without radiologist interpretation were excluded. Ultimately, 956 non-contrast brain CT scans were included in this study.

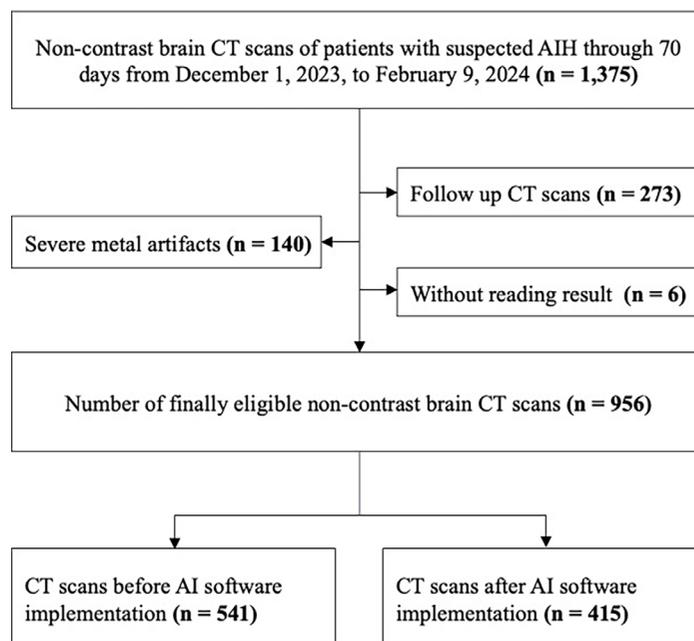
To distinguish between study periods before and after AI software implementation, the boundary date was set as January 5, 2024, the date of implementation. Consequently, the pre-AI period was defined as the 35 days from December 1, 2023, to January 4, 2024, whereas the post-AI period covered the following 35 days from January 5 to February 9, 2024. Of the 956 brain CT scans, 541 were acquired during the pre-AI period, and the remaining 415 during the post-AI period (Figure 1).

### Computed tomography scanning protocol

CT scans were performed using one of two CT machines at the institution. Machine A was a 128-slice single-source CT scanner (SOMATOM Edge, Siemens Healthineers, Forchheim, Germany) with a tube potential of 70–140 kVp and 20–800 mA; machine B was a dual-source CT scanner (SOMATOM Force, Siemens Healthineers, Germany) with

### Main points

- A commercially available computed tomography-based artificial intelligence (AI) software was developed to ease the growing burden on radiologists to promptly diagnose acute intracranial hemorrhage (AIH).
- Evaluating AI software in a real-world clinical setting is essential for practical use.
- The implementation of this AI software considerably optimized radiologists' prioritization of reading order and enabled earlier reporting of AIH cases without compromising performance.
- The optimized workflow by the AI software integration is expected to improve the prompt and efficient treatment of patients with AIH.



**Figure 1.** Flowchart of study enrollment. AI, artificial intelligence; AIH, acute intracranial hemorrhage; CT, computed tomography.

a variable tube potential of 70–150 kVp and 20–1300 mA. The acquisition parameters were as follows: slice thickness, 4 mm without gap; rotation time, 1.0 s; pitch, 1; automatic tube voltage modulation (CARE kV, Siemens Healthineers, Germany) using a reference of 120 kV; automatic tube current selection (CAREdose 4D, Siemens Healthineers, Germany) using a reference of 250 mAs; and collimation of 128 × 0.6 for machine A and 192 × 0.6 for machine B.

### Artificial intelligence software development

The commercially available CT-based AI software for detecting AIH (HyperInsight - ICH, version 2.0.1, Purple AI Inc., Korea) used in this study was developed using deep learning algorithms trained on 28,351 slices from 2,010 patients with AIH and 1,000 normal participants. The AIH detection process employed a joint convolutional and recurrent neural network-based sequence module that provided AIH probability scores (ranging from 0 to 100) on both a patient-wise and slice-wise basis. It also generated anomalies for patients with AIH by subtracting original CT images from restored images and post-processing them using unsupervised training on normal datasets. AI-assisted brain CT images showing AIH locations and scores were displayed to the radiologists on the PACS viewer alongside the original images.<sup>18</sup>

### Ground truth for acute intracranial hemorrhage

To establish the ground truth for AIH, 2 board-certified neuroradiologists (S.W.O. and H.Y.L., with 17 and 19 years of experience in brain imaging, respectively) independently reviewed the same set of 956 non-contrast brain CT scans. The neuroradiologists diagnosed AIH based solely on CT findings and were blinded to patients' clinical information, previous reading results, and follow-up imaging. In cases of disagreement, the ground truth was determined by consensus, referring to other available imaging modalities.

### Radiologists' computed tomography interpretation

Two board-certified general radiologists (H.B. and H.S., each with 10 years of experience in brain imaging without fellowship training in neuroradiology) routinely interpreted the enrolled non-contrast brain CT scans as part of clinical practice. These radiologists were blinded and unaware of the study's purpose and design throughout the entire study period. Therefore, they could

freely refer to patients' clinical information and other available studies using the institution's PACS (ZeTTA PACS, version 1.0.0.42.10, TaeYoung Soft, Korea). Prior to AI software implementation, the two radiologists received brief training in using the software from a board-certified neuroradiologist (J.K.) for 1 day. The radiologists required minimal learning time with the AI software, as the probability scores were intuitively presented within the existing worklist interface. After implementation, the AIH probability scores generated by the software were integrated into the PACS worklist, allowing the radiologists to determine the reading order based on the scores. Figure 2 exemplifies the worklists before and after implementation. During the entire study period, CT scan completion time and the radiologists' final report time were automatically recorded on the PACS server of our institution.

### Definition of the early diagnosis rate

Since early diagnosis of AIH is crucial for improving patient outcomes,<sup>1,4</sup> the early diagnosis rate for AIH cases was defined to assess the potential effectiveness of changes to the reading order. The first quartile of the radiologists' reading order was chosen as the threshold for defining early diagnosis, because the first quartile is commonly used to identify the highest-priority or most urgent cases in general medical practice.<sup>25,26</sup> By using the first quartile of reads, the aim was to assess the effectiveness of the prioritization by the AI software. The equation for the early diagnosis rate was defined as follows:

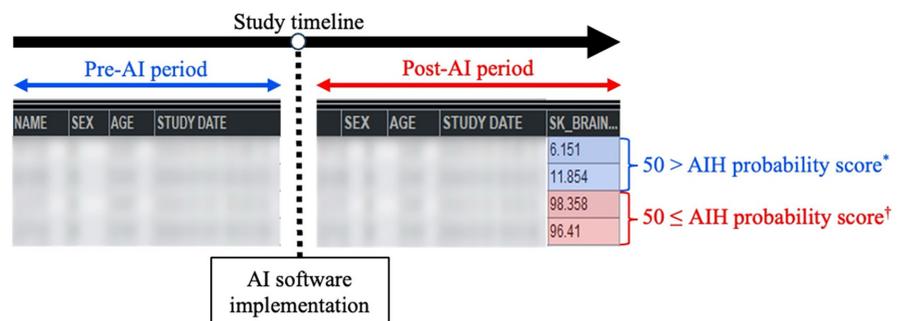
$$\text{Early Diagnosis Rate} = \frac{\text{AIH cases read rapidly within the first quartile by radiologists}}{\text{Total AIH cases}} \quad (1)$$

### Statistical analysis

The sample size of the case group was calculated based on a significance level of 0.05, a statistical power of 0.8, a specificity of 0.90 from a previous meta-analysis, and a specificity of 0.984 from prior validation research, with a dropout rate of 10%.<sup>13,17</sup> The determined sample size for the study was 202 cases. Due to its explanatory nature, the sample size for the daily analysis was determined based on previous studies,<sup>27</sup> and a minimum of 1 month was selected for each period before and after AI software implementation. The stand-alone performance of the AI software after implementation was evaluated using the area under the receiver operating characteristic curve (AUC), accuracy, sensitivity, and specificity.

First, a simple comparison of the radiologists' absolute TAT (the time gap between CT scan completion and the radiologists' final report) was conducted as a preliminary study. The TAT of cases with and without AIH between the pre-and post-AI periods was compared using an independent t-test, following the Shapiro-Wilk test for normality. This preliminary comparison aimed to explore the feasibility of conducting daily comparisons and to avoid bias arising from TAT comparisons.

Furthermore, to evaluate the impact of AI software on the radiologists' daily diagnostic performance for AIH, their accuracy, sensitivity, and specificity were calculated in both pre-and post-AI periods and compared between the two periods. Moreover, the impact of false negative and false positive cases generated by the AI software on radiologists' decisions was assessed in an additional sub-analysis.



**Figure 2.** This figure provides examples of the worklists used in the study. The left worklist during the pre-AI period shows the radiologists' routine worklists before AI software implementation. By contrast, in the SK\_BRAIN column, the AIH probability scores generated by the software were added to the worklist during the post-AI period. Consequently, the radiologists could use the score to predict that the cases with a blue background (\*) were less likely to exhibit AIH, whereas the cases with a red background were more likely to exhibit AIH (†) after the implementation. AI, artificial intelligence; AIH, acute intracranial hemorrhage.

Lastly, the impact of the AI software on the radiologists' workflow was evaluated. The ordinal correlation between the chronological order of CT scans and the radiologists' reading order was measured using Kendall's Tau in both pre- and post-AI periods. These rank correlation coefficients were compared between the two periods. In addition, the modified reading order was evaluated to confirm whether it appropriately prioritized the rapid reading of AIH. For this evaluation, the daily early diagnosis rate for AIH cases and the median reading order of AIH were calculated in both pre- and post-AI periods and compared between the two periods.

Mean daily diagnostic performance; Kendall's Tau; early diagnosis rate for AIH cases; median reading order of AIH; and baseline characteristics including age, gender proportion, AIH incidence, Glasgow Coma Scale scores, and modified Rankin scale scores between the pre- and post-AI periods were compared using independent t-tests or Mann-Whitney U tests following the Shapiro-Wilk test. A visual summary of the comparison analyses is presented in Figure 3.

Continuous variables were described as means with 95% confidence intervals (CIs) using bootstrapping, and ordinal variables were described as medians with ranges from the 25<sup>th</sup> percentile (Q1) to the 75<sup>th</sup> percentile (Q3). The statistical software MedCalc (version 23.2.1, MedCalc Software Ltd, USA) was used for statistical analysis. A *P* value less than 0.05 was considered statistically significant.

## Results

### Patient characteristics

A total of 956 initial CT scans from 956 patients were included. Of these, 541 and 415

CT scans were acquired during the pre- and post-AI periods, respectively. The mean age of the total patient cohort was 63.14 years  $\pm$  18.41 (standard deviation), the proportion of male participants was 45%, and the incidence of AIH was 13%. There was no significant difference in median age [pre-AI period: 67 years (51–77); post-AI period: 67 (52–78); *P* = 0.558], number of male patients [pre-AI period: 246 (45%); post-AI period: 201 (48%); *P* = 0.363], AIH cases [pre-AI period: 72 (13%); post-AI period: 50 (12%); *P* = 0.681], median Glasgow Coma Scale score [pre-AI period: 15 (15–15); post-AI period: 15 (15–15); *P* = 0.831], and modified Rankin scale scores [pre-AI period: 0 (0–0); post-AI period: 0 (0–0); *P* = 0.295] before and after AI implementation. The number of daily CT scans [pre-AI period: 12 (7.25–17.75); post-AI period: 12 (10–19.75); *P* = 0.256] and daily AIH cases [pre-AI period: 1 (0–1.75); post-AI period: 2 (1–3); *P* = 0.063] were not significantly different. These results are summarized in Table 1.

### Preliminary comparison of turnaround time

In the preliminary study, the mean TAT significantly decreased (from 1,610 min to 1,145 min, *P* < 0.001) after AI software imple-

mentation. When analyzed by cases with and without AIH, TAT significantly decreased in both cases with AIH (from 1,452 min to 870 min, *P* < 0.001) and without AIH (from 2,084 min to 1,184 min, *P* < 0.001) after AI software implementation. These preliminary results are illustrated in Figure 4.

### Stand-alone performance of the artificial intelligence software

The prevalence of AIH in the post-AI period was 12%. After AI software implementation, the AUC for the standalone AI software was 0.99 (95% CI, 0.98–0.99) in detecting AIH. The accuracy, sensitivity, and specificity were 0.98 (95% CI, 0.97–0.99), 0.96 (95% CI, 0.86–0.99), and 0.99 (95% CI, 0.97–0.99), respectively, using a probability score cut-off of 50% for detecting AIH.

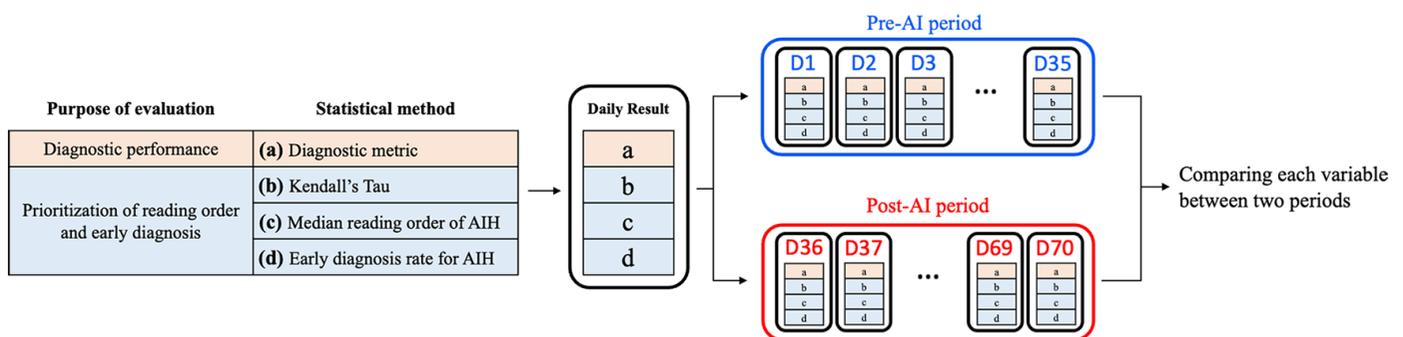
### Diagnostic performance of radiologists

The radiologists' daily accuracy [from 0.99 (95% CI, 0.99–1.00) to 0.99 (95% CI, 0.98–1.00), *P* = 0.343], sensitivity [from 1.00 (95% CI, 0.99–1.00) to 1.00 (95% CI, 0.99–1.00), *P* = 0.859], and specificity [from 1.00 (95% CI, 0.99–1.00) to 0.99 (95% CI, 0.97–1.00), *P* = 0.252] for detecting AIH were not significant.

**Table 1.** Baseline characteristics of patients

Baseline characteristics	Pre-AI period (n = 541)	Post-AI period (n = 415)	<i>P</i> value
Age (years)*	67 (51–77)	67 (52–78)	0.558
Number of male patients (%)	246 (45)	201 (48)	0.363
Number of AIH cases (%)	72 (13)	50 (12)	0.681
Glasgow Coma Scale score*	15 (15–15)	15 (15–15)	0.831
Modified Rankin Scale score*	0 (0–0)	0 (0–0)	0.295
Number of daily CT scans*	12 (7.25–17.75)	12 (10–19.75)	0.256
Number of daily AIH cases*	1 (0–1.75)	2 (1–3)	0.063

\*The Mann-Whitney U test was used, and the values are presented as the median with the range between Q1 and Q3. AI, artificial intelligence; AIH, acute intracranial hemorrhage; CT, computed tomography.



**Figure 3.** This figure presents a schematic representation of the statistical analyses used in this study. To assess the AI software's impact on the radiologists' diagnostic performance (a) in AIH detection, performance was compared before and after implementation. To assess the impact of the software on the radiologists' workflow, Kendall's Tau (b) was used to compare the correlation between the daily chronological order of the CT scan and the radiologists' reading order before and after the implementation. The median reading order of AIH (c) and the early diagnosis rate for AIH (d) (defined as the proportion of AIH cases read rapidly within the top quarter by radiologists) were compared before and after implementation. AI, artificial intelligence; AIH, acute intracranial hemorrhage; CT, computed tomography.

ly different after AI software implementation. These results are summarized in Table 2.

In an additional sub-analysis of false negative and false positive cases, there were two false negative and four false positive cases generated by the AI software. However, the radiologists' diagnoses and the ground truth for AIH were entirely identical even in these cases. Examples of cases with and without AIH integrated with the AI software are illustrated in Figure 5.

### Prioritization of reading order and early diagnosis

The daily correlation between the chronological order of CT scans and the radiologists' reading order significantly decreased after AI software implementation [Kendall's Tau: from 0.61 (95% CI, 0.48–0.73) to 0.01 (95% CI, 0.00–0.26),  $P < 0.001$ ]. The radiologists' daily early diagnosis rate of AIH significantly increased after AI software implementation [from 0.50 (0.23–1.00) to 1.00 (0.55–1.00),  $P = 0.014$ ]. Furthermore, the radiologists' daily median reading order for AIH cases significantly decreased after AI software implementation [from 7.25 (3–11.75) to 1.5 (1–3),  $P < 0.001$ ]. These results are summarized in Table 2 and illustrated in Figure 6.

## Discussion

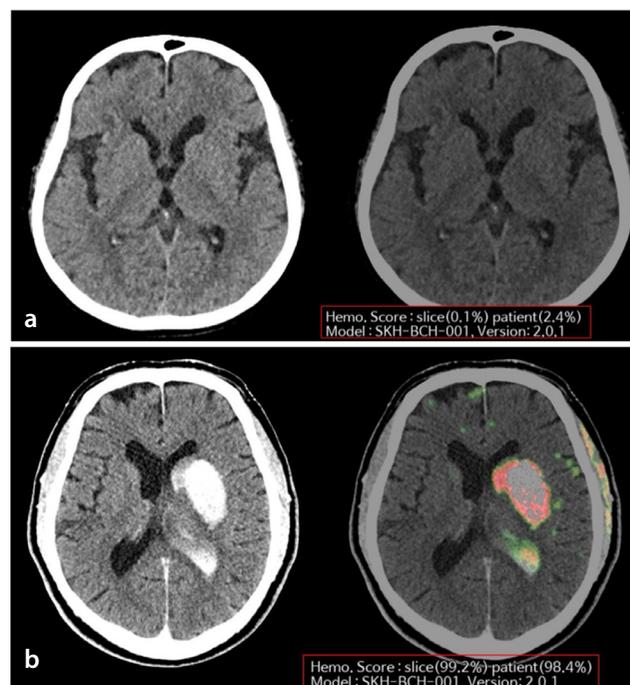
This study aimed to assess the impact of a commercially available CT-based AI software for AIH detection on radiologists' diagnostic performance and workflow. The software greatly optimized radiologists' reading prioritization and enabled them to read AIH cases more rapidly in daily practice. Furthermore, the AI software did not compromise the radiologists' diagnostic performance for detecting AIH, even in cases where the AI generated false positives or false negatives.

Regarding the radiologists' diagnostic performance for AIH, the impact of the AI software was negligible, and the radiologists were not influenced by the false negative or false positive results generated by the software. Several factors may explain this finding. First, the study design played a role. In this observational study, the readers had access to patient information and other examinations as part of routine clinical practice, unlike previous validation studies with controlled conditions where readers lacked clinical context.<sup>17</sup> Additionally, the diagnostic accuracy of board-certified radiologists for AIH is known to be particularly high in routine clinical settings.<sup>1–3</sup> Therefore, it is not surprising that the radiologists in

this study—being board-certified and experienced in diagnosing AIH—maintained high performance. Notably, the minor changes in accuracy and specificity may indicate effective management of false positives by the AI software. In other words, potential false positives generated by the AI were either easily recognized or efficiently disregarded, thereby not compromising diagnostic outcomes. Consequently, our findings suggest that the AI software's impact on detection performance may be negligible—or at least not detrimental—when radiologists interpret images under routine conditions or already possess sufficient diagnostic expertise.<sup>28,29</sup>

To evaluate whether the AI software could influence the radiologists' actual reading or-

der, we compared the correlation between the chronological order of CT scans and the radiologists' reading order before and after AI software implementation. Before the implementation, there was a high correlation between the two, suggesting that radiologists typically interpreted CT scans using a traditional stat or FIFO prioritization system. However, after implementation, a considerable dissociation between the two orders was observed, along with an increased early diagnosis rate of AIH. This suggests that the integrated AI software substantially altered the radiologists' reading order and facilitated prioritization of CT scans with AIH over those without. This shift in prioritization occurred because radiologists could estimate the ur-



**Figure 4.** In each figure, the left image shows a non-contrast brain CT scan, whereas the right image shows an overlaid heatmap of AIH generated by the AI software. The red box below depicts the AIH probability score (hemo. score) for both the slice- and patient-wise levels, as well as the model and version information of the AI software. (a) AIH is not visible on the CT scan, resulting in an AIH probability score of less than 50%, with no heatmap. (b) There is AIH in the left basal ganglia, extending to the left lateral ventricle, resulting in a probability score of over 50% with a visible heatmap. AI, artificial intelligence; AIH, acute intracranial hemorrhage; CT, computed tomography.

**Table 2.** Radiologists' diagnostic performance, prioritization of reading order, and early diagnosis between the pre- and post-AI periods

Variables	Pre-AI period (35 days)	Post-AI period (35 days)	P value
Accuracy <sup>†</sup>	0.99 (0.99–1.00)	0.99 (0.98–1.00)	0.343
Sensitivity <sup>†</sup>	1.00 (0.99–1.00)	1.00 (0.99–1.00)	0.859
Specificity <sup>†</sup>	1.00 (0.99–1.00)	0.99 (0.97–1.00)	0.252
Kendall's Tau <sup>†</sup>	0.61 (0.48–0.73)	0.01 (0.00–0.26)	<0.001*
Early diagnosis rate <sup>‡</sup>	0.50 (0.23–1.00)	1.00 (0.55–1.00)	0.014*
Median reading order <sup>‡</sup>	7.25 (3.00–10.75)	1.50 (1.00–3.00)	<0.001*

\* $P < 0.05$ , statistical significance. <sup>†</sup>An independent t-test was used, and the values are presented as the mean with 95% CI. <sup>‡</sup>The Mann-Whitney U test was used, and the values are presented as the median with the range between Q1 and Q3. AI, artificial intelligence; CI, confidence interval.

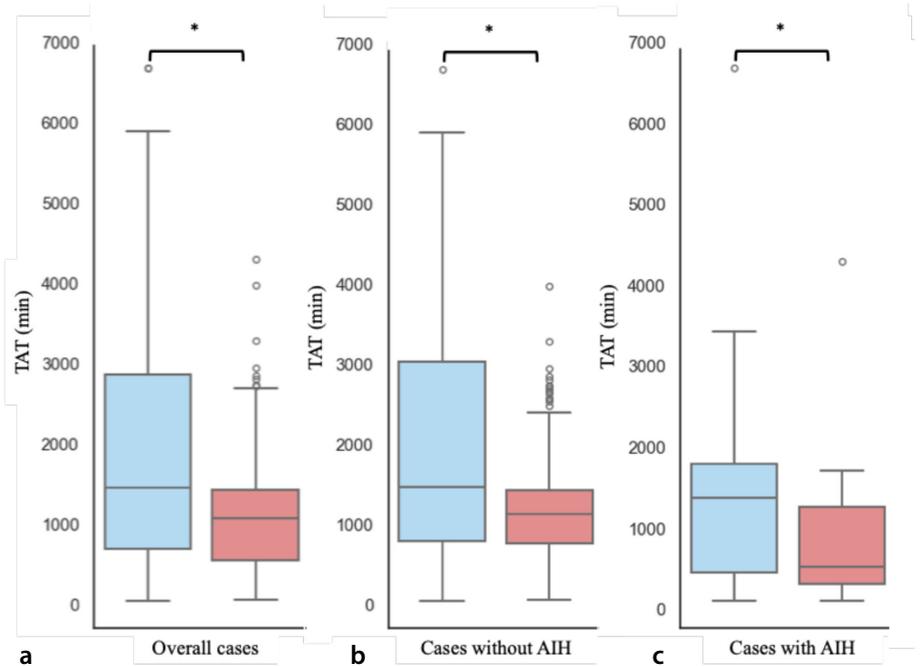
gency of AIH cases by referring to the AIH probability score before opening a CT scan from their worklist. This predictability led to a remarkable increase in early diagnosis. After implementation, the median reading order of AIH cases considerably decreased, and the early diagnosis rate for AIH cases increased substantially. These changes signify that the radiologists' workflow was prioritized and optimized to allow for more rapid interpretation of AIH cases. Considering that non-contrast brain CT is the first-line approach for AIH, these improvements brought by the AI software may enhance not only the promptness but also the efficiency of clinical diagnosis and treatment for patients with suspected AIH.<sup>1-4,6,24</sup>

In terms of patient characteristics, the modified Rankin scale scores were not considerably different after AI software implementation. However, these findings should be interpreted with caution. Because the primary objective of this study was to evaluate the impact of AI integration on radiologists' workflow, the AI software was not utilized by physicians in clinical decision-making. Moreover, functional outcomes are influenced by a wide range of clinical variables, including age, neurological status, comorbidities, and treatment delays.<sup>2-4,16</sup> None of these factors were adjusted for in our analysis, as this was beyond the scope of the study. Therefore, the lack of observed improvement in functional outcomes does not imply that the AI software lacks clinical value. On the contrary, considering our findings demonstrating enhanced reading prioritization by AI and previous research indicating the greatest benefits of AI when used by clinicians,<sup>17</sup> it can be inferred that AI contributes to efficiency and potentially improves patient care in clinical environments. Consequently, this study remains important as it establishes a foundation for the broader adoption of AI in clinical practice.

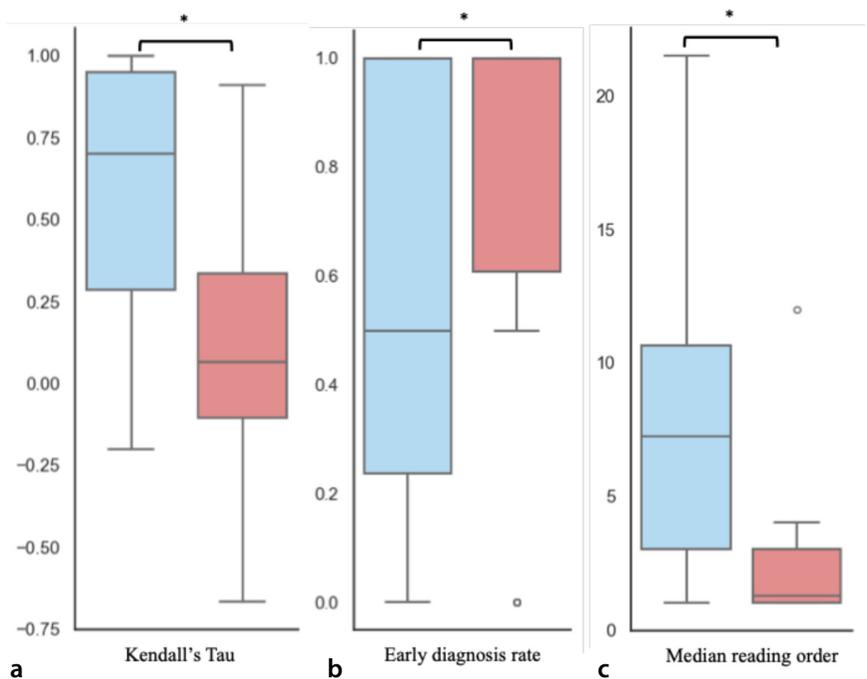
In this study, we conducted an ordinal comparison on a daily basis rather than a simple TAT comparison between the pre-AI and post-AI periods, as the mean TAT for both cases with and without AIH had already decreased substantially in the preliminary study. Radiologists' TAT can be affected by numerous factors, including routine tasks, working days, or other unexpected circumstances,<sup>7-11</sup> and the radiologists in this study—who interpreted various imaging modalities across different body parts—may have been similarly influenced.<sup>19,20</sup> Therefore, our daily ordinal comparison of radiologists' reading order more accurately reflected their work-

flow in a routine real-world clinical setting than a simple TAT comparison. As a result, we mitigated potential bias and gained clearer insights into radiologists' workflow.

This study had several limitations. First, its retrospective observational design may have introduced uncontrolled bias that could have affected our results. Second, the findings were based on data from a single insti-



**Figure 5.** This figure illustrates box and whisker plot charts comparing TAT between the pre-AI (blue box) and post-AI (red box) periods. In the overall case (a), the mean TAT considerably decreased after AI software implementation. The mean TATs of cases without (b) and with AIH (c) in the post-AI period were considerably shorter than those in the pre-AI period. An asterisk (\*) indicates a statistically significant difference. AI, artificial intelligence; AIH, acute intracranial hemorrhage; TAT, turnaround time.



**Figure 6.** This figure illustrates box and whisker plot charts comparing variables between the pre-AI (blue box) and post-AI (red box) periods. (a) Kendall's Tau between the chronological order and (b) the radiologists' reading order considerably decreased, whereas the early diagnosis rate for AIH considerably increased after AI software implementation. (c) The median reading order of AIH decreased substantially after implementation. An asterisk (\*) indicates a statistically significant difference. AI, artificial intelligence; AIH, acute intracranial hemorrhage.

tution using machines from a single vendor, which may limit the generalizability of the study. Additionally, radiologists' experience levels, institutional CT workflow protocols, and the availability of technical support may vary greatly across centers, potentially influencing both diagnostic performance and the impact of AI-driven prioritization.<sup>30</sup> Third, our statistical analysis of daily comparisons for radiologists' performance and workflow, while logically sound, may have unpredictably weakened statistical power by reducing the sample size from hundreds to dozens. To maintain statistical robustness without sacrificing temporal granularity, future research could employ rolling averages, time-series models that account for intraday variability, or extend the study period. Finally, this comparison study focused solely on the daily impact of AI software assistance on AIH detection within the radiologists' workflow and did not assess broader real-world challenges. For instance, integrating AI into clinical workflows requires substantial computational resources and careful implementation planning. Therefore, additional prospective multicenter trials involving multiple vendors, a larger reader cohort, and diverse clinical settings are needed to mitigate potential selection bias and improve generalizability.<sup>30,31</sup>

In conclusion, the integration of CT-based AI software for detecting AIH considerably enhanced the prioritization of radiologists' reading order and accelerated their interpretation of AIH cases while maintaining diagnostic performance by optimizing workflows in real-world clinical settings. Consequently, with the increasing number of CT scans and the growing demands placed on radiologists, AI software is expected to improve workflow efficiency and support the prompt and effective treatment of patients with AIH.

### Acknowledgement

We sincerely thank two contributors - Dr. Hosuk Song and Dr. Hokyun Byun for their valuable contributions and support in the development of this manuscript.

### Footnotes

### Conflict of interest disclosure

The authors declared no conflicts of interest.

### Funding

This work was supported by the National Research Foundation of Korea funded by the Korean government (MSIT) [RS-2023-

00208409]; and the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea [HR22C173401].

## References

1. Waqas M, Vakharia K, Munich SA, et al Initial emergency room triage of acute ischemic stroke. *Neurosurgery*. 2019;85(suppl\_1):S38-S46. [\[Crossref\]](#)
2. van Asch CJ, Luitse MJ, Rinkel GJ, van der Tweel I, Algra A, Klijn CJ. Incidence, case fatality, and functional outcome of intracerebral haemorrhage over time, according to age, sex, and ethnic origin: a systematic review and meta-analysis. *Lancet Neurol*. 2010;9(2):167-176. [\[Crossref\]](#)
3. Broderick JP, Brott TG, Duldner JE, Tomsick T, Huster G. Volume of intracerebral hemorrhage. A powerful and easy-to-use predictor of 30-day mortality. *Stroke*. 1993;24(7):987-993. [\[Crossref\]](#)
4. Forman R, Slota K, Ahmad F, et al. Intracerebral hemorrhage outcomes in the very elderly. *J Stroke Cerebrovasc Dis*. 2020;29(5):104695. [\[Crossref\]](#)
5. National Council on Radiation Protection and Measurements. Scientific Committee 6-2 on Radiation Exposure of the U.S. Population. *Ionizing Radiation Exposure of the Population of the United States*. Bethesda (MD): National Council on Radiation Protection and Measurements; 2009. (Report No.: 160). [\[Crossref\]](#)
6. Winder M, Owczarek AJ, Chudek J, Pilch-Kowalczyk J, Baron J. Are we overdoing it? Changes in diagnostic imaging workload during the years 2010-2020 including the impact of the SARS-CoV-2 pandemic. *Healthcare (Basel)*. 2021;9(11):1557. [\[Crossref\]](#)
7. Kansagra AP, Liu K, Yu JP. Disruption of radiologist workflow. *Curr Probl Diagn Radiol*. 2016;45(2):101-106. [\[Crossref\]](#)
8. Mamlouk MD, Saket RR, Hess CP, Dillon WP. Adding value in radiology: establishing a designated quality control radiologist in daily workflow. *J Am Coll Radiol*. 2015;12(8):838-841. [\[Crossref\]](#)
9. Kotter E, Ranschaert E. Challenges and solutions for introducing artificial intelligence (AI) in daily clinical workflow. *Eur Radiol*. 2021;31(1):5-7. [\[Crossref\]](#)
10. Balint BJ, Steenburg SD, Lin H, Shen C, Steele JL, Gunderman RB. Do telephone call interruptions have an impact on radiology resident diagnostic accuracy? *Acad Radiol*. 2014;21(12):1623-1628. [\[Crossref\]](#)
11. Yu JP, Kansagra AP, Mongan J. The radiologist's workflow environment: evaluation of disruptors and potential implications. *J Am Coll Radiol*. 2014;11(6):589-93. [\[Crossref\]](#)
12. Halsted MJ, Froehle CM. Design, implementation, and assessment of a radiology workflow management system. *AJR Am J Roentgenol*. 2008;191(2):321-7. [\[Crossref\]](#)
13. Agarwal S, Wood D, Grzeda M, et al. Systematic review of artificial intelligence for abnormality detection in high-volume neuroimaging and subgroup meta-analysis for intracranial hemorrhage detection. *Clin Neuroradiol*. 2023;33(4):943-956. [\[Crossref\]](#)
14. Mouridsen K, Thurner P, Zaharchuk G. Artificial intelligence applications in stroke. *Stroke*. 2020;51(8):2573-2579. [\[Crossref\]](#)
15. Segato A, Marzullo A, Calimeri F, De Momi E. Artificial intelligence for brain diseases: a systematic review. *APL Bioeng*. 2020;4(4):041503. [\[Crossref\]](#)
16. de Havenon A, Tirschwell DL, Heitsch L, et al. Variability of the modified Rankin scale score between day 90 and 1 year after ischemic stroke. *Neurol Clin Pract*. 2021;11(3):e239-e244. [\[Crossref\]](#)
17. Yun TJ, Choi JW, Han M, et al. Deep learning based automatic detection algorithm for acute intracranial haemorrhage: a pivotal randomized clinical trial. *NPJ Digit Med*. 2023;6(1):61. [\[Crossref\]](#)
18. Kim J, Oh SW, Lee HY, et al. Assessment of deep learning-based triage application for acute ischemic stroke on brain MRI in the ER. *Acad Radiol*. 2024;31(11):4621-4628. [\[Crossref\]](#)
19. Wesp W. Using STAT properly. *Radiol Manage*. 2006;28(1):26-30; quiz 31-33. [\[Crossref\]](#)
20. Gaskin CM, Patrie JT, Hanshew MD, Boatman DM, McWey RP. Impact of a reading priority scoring system on the prioritization of examination interpretations. *AJR Am J Roentgenol*. 2016;206(5):1031-1039. [\[Crossref\]](#)
21. Kotovich D, Twig G, Itsekson-Hayosh Z, et al. The impact on clinical outcomes after 1 year of implementation of an artificial intelligence solution for the detection of intracranial hemorrhage. *Int J Emerg Med*. 2023;16(1):50. [\[Crossref\]](#)
22. Zia A, Fletcher C, Bigwood S, et al. Retrospective analysis and prospective validation of an AI-based software for intracranial haemorrhage detection at a high-volume trauma centre. *Sci Rep*. 2022;12(1):19885. [\[Crossref\]](#)
23. Baltruschat I, Steinmeister L, Nickisch H, et al. Smart chest X-ray worklist prioritization using artificial intelligence: a clinical workflow simulation. *Eur Radiol*. 2021;31(6):3837-3845. [\[Crossref\]](#)
24. McWey RP, Hanshew MD, Patrie JT, Boatman DM, Gaskin CM. Impact of a four-point order-priority score on imaging examination performance times. *J Am Coll Radiol*. 2016;13(3):286-95.e5. [\[Crossref\]](#)
25. Maltby J, Williams G, McGarry J, Day L. Research methods for nursing and healthcare. Routledge; 2014. [\[Crossref\]](#)

26. Kim B, Romeijn S, van Buchem M, Mehrizi MHR, Grootjans W. A holistic approach to implementing artificial intelligence in radiology. *Insights Imaging*. 2024;15(1):22. [\[Crossref\]](#)
27. Savage CH, Tanwar M, Elkassem AA, et al. Prospective evaluation of artificial intelligence triage of intracranial hemorrhage on noncontrast head CT examinations. *AJR Am J Roentgenol*. 2024;223(5):e2431639. [\[Crossref\]](#)
28. Yang HK, Ko Y, Lee MH, et al. Initial performance of radiologists and radiology residents in interpreting low-dose (2-mSv) appendiceal CT. Erratum in: *AJR Am J Roentgenol*. 2016;206(4):901. [\[Crossref\]](#)
29. Labus S, Altmann MM, Huisman H, et al. A concurrent, deep learning-based computer-aided detection system for prostate multiparametric MRI: a performance study involving experienced and less-experienced radiologists. *Eur Radiol*. 2023;33(1):64-76. [\[Crossref\]](#)
30. Drukker K, Chen W, Gichoya J, et al. Toward fairness in artificial intelligence for medical image analysis: identification and mitigation of potential biases in the roadmap from data collection to model deployment. *J Med Imaging (Bellingham)*. 2023;10(6):061104. [\[Crossref\]](#)
31. Panayides AS, Amini A, Filipovic ND, et al. AI in medical imaging informatics: current challenges and future directions. *IEEE J Biomed Health Inform*. 2020;24(7):1837-1857. [\[Crossref\]](#)